(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 641 861 B1

B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
09.04.2003 Bulletin 2003/15

(51) Int Cl.[7]: C12N 15/62, C07K 14/16, C07K 14/47, C07K 19/00, C12N 15/70, C12N 1/21

(21) Application number: 94109676.0

(22) Date of filing: 23.06.1994

(54) **Process for producing hydrophobic polypeptides, proteins or peptides**

Verfahren zur Herstellung von hydrophoben Polypeptiden, Proteinen und Peptiden

Procédé de production de polypeptides, proteines et peptides hydrophobiques

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: 06.07.1993 EP 93110755

(43) Date of publication of application:
08.03.1995 Bulletin 1995/10

(60) Divisional application:
03000052.5

(73) Proprietor: F. HOFFMANN-LA ROCHE AG
4002 Basel (CH)

(72) Inventors:
- Döbeli, Heinz
CH-4417 Ziefen (CH)
- Draeger, Nicholas
CH-4056 Basle (CH)
- Huber Trottman, Gerda
CH-4247 Grindel (CH)
- Jakob, Peter
CH-4153 Reinach (CH)
- Stüber, Dietrich
D-79639 Grenzach-Wyhlen (DE)

(56) References cited:
EP-A- 0 289 110
EP-A- 0 353 823
WO-A-92/03474
WO-A-93/04194

- NATURE vol. 325, 19 February 1987, pages 733 - 736, XP002023114 J. KANG ET AL: "The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor"
- VACCINE, vol. 5, no. 2, June 1987, pages 90-101, XP002023125 P. BARR ET AL: "Antigenicity and immunogenicity of domains of the human immunodeficiency virus (HIV) envelope polypeptide expressed in the yeast Saccharomyces cerevisiae"
- JOURNAL OF MEDICAL VIROLOGY, vol. 23, no. 1, September 1987, pages 1-9, XP002023126 F. CHIODI ET AL: "Site-directed ELISA with synthetic peptides representing the HIV transmembrane glycoprotein"

EP 0 641 861 B1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The possibilities of preparing hybrid genes by gene technology open up new routes for the working-up of recombinant proteins. By linking the coding gene sequence of a desired protein with the coding gene sequence of a protein fragment having a high affinity for a ligand (affinity peptide) it is possible to purify desired recombinant proteins in the form of fusion proteins in one step using the affinity peptide. By site-directed mutagenesis it is also possible to introduce specific chemical or enzymatic cleavage sites at the point of linkage of the affinity peptide and the desired recombinant protein, so that after the purification of the fusion protein by means of a suitable affinity resin the desired recombinant protein can be recovered by chemical or enzymatic cleavage.

[0002] However, recovering of the desired recombinant protein may turn out to be extremely difficult when the desired recombinant protein also contains such chemical or enzymatic cleavage sites in its amino acid sequence. In such cases, the desired recombinant protein ends up being rapidly degraded.

[0003] In order to inhibit this degradation, the present invention provides fusion proteins and processes which allow the selective cleavage at a specific chemical or enzymatic cleavage site without affecting the desired recombinant protein. The methods of the present invention are specifically applicable for production of hydrophobic polypeptides, proteins or peptides.

[0004] In more detail, the present invention is concerned with fusion proteins of the formula:

$$A - B - C$$

wherein A is a bulky hydrophilic peptide, B is a selected cleavage site, and C is a desired hydrophobic polypeptide, protein or peptide.

[0005] The term "bulky hydrophilic peptide" which is used in connection with the fusion proteins in accordance with the invention relates to hydrophilic peptides which are characterized by their size and the content of hydrophilic amino acids giving rise to a well structured domain.

[0006] The bulky hydrophilic peptide A of the fusion proteins in accordance with the invention serves a dual function: a) to facilitate high expression of the fusion proteins and b) to expose the cleavage site C to the mobile phase on a hydrophobic matrix column. Preferred bulky hydrophilic peptides of the fusion proteins in accordance with the invention are those with the peptide sequence of the formula

$$(NANP)\ x$$

wherein x is 10 - 40 and 19 being most preferred.

[0007] As selected cleavage sites there come into consideration chemical or enzymatic cleavage sites. As suitable selected enzymatic cleavage sites there come into consideration the amino acid sequences -$(Asp)_n$-Lys-, wherein n signifies 2, 3 or 4, or -Ile-Glu-Gly-Arg- which can be specifically recognized by the proteases enterokinase and coagulation factor X_a, respectively, an arginine residue or a lysine residue cleaved by trypsin, a lysine residue cleaved by lysyl endopeptidase or a glutamine residue cleaved by V8 protease. As suitable selected chemical cleavage sites there come into consideration tryptophan residues cleaved by 3-bromo-3-methyl-2- (2-nitrophenylmercapto)-3H-indole, cystein residues cleaved by 2-nitroso-5-thiocyano benzoic acid, the amino acid dipeptides Asp-Pro or Asn-Gly which can be cleaved by acid and hydroxylamine, respectively, and preferably, a methionine residue which is specifically cleaved by cyanogen bromide (CNBr).

[0008] The term "hydrophobic polypeptide, protein or peptide" which is used in connection with the fusion proteins in accordance with the invention relates to hydrophobic polypeptides, proteins or peptides which elute from reversed phase HPLC columns at concentrations between 30 and 60%, preferably around 40% of organic solvents in aqueous buffer, e.g. at a concentration of higher than 40% ethanol in aqueous buffer.

[0009] As hydrophobic polypeptides, proteins or peptides there come into consideration, for example, surface antigens, lymphokine receptors, HIV-1 and HIV-2 envelope and structure proteins, hepatitis C envelope and structure proteins or any peptide with membrane anchor sequences. A preferred hydrophobic peptide is the peptide having the sequence:

1 35
RILAVERYLKDQQLLGIWGCSGKLICTTAVPWNAS (Seq ID No:11)

The bulky hydrophilic peptides and the selected cleavage sites of the fusion proteins in accordance with the invention can be linked either to the amino terminal amino acid or to the carboxy terminal amino acid of the hydrophobic polypeptide, protein or peptide.

[0010] The fusion proteins in accordance with the present invention can optionally contain specific sequences that preferably bind to an affinity carrier material. Examples of such sequences are sequences containing at least two adjacent histidine residues (see in this respect European Patent Application Publication No. 282 042). Such sequences bind selectively to nitrilotriacetic acid nickel chelate resins (Hochuli and Döbeli, Biol.Chem. Hoppe-Seyler 368, 748 (1987); European Patent No. 253 303). Fusion proteins of the present invention which contain such a specific sequence can, therefore, be separated selectively from the remaining polypeptides. The specific sequence can be linked either to the amino acid sequence of the bulky hydrophilic peptide or the amino acid sequence of the hydrophobic polypeptide, protein or peptide.

[0011] The present invention is also concerned with genes which code for these fusion proteins, expression vectors which contain these genes, microorganisms transformed with these expression vectors as well as a process for the preparation of said genes, expression vectors and transformed microorganisms.

[0012] The preparation of the fusion proteins in accordance with the invention can be effected according to methods of recombinant DNA technology which are described in the literature. Preferably, a nucleotide sequence coding for the desired hydrophobic polypeptide, protein or peptide is firstly synthesized and this is then linked with a nucleotide sequence coding for the bulky hydrophilic peptide and the selected cleavage site.

[0013] The incorporation of the thus-obtained hybrid gene in expression vectors is also effected in a manner known per se. In this context reference can be made to the textbooks of Maniatis et al. ("Molecular Cloning", Cold Spring Harbor Laboratory, 1982) and Sambrook et al. ("Molecular Cloning-A Laboratory Manual", 2nd. ed., Cold Spring Harbor Laboratory, 1989).

[0014] The methods for the expression of the fusion proteins in accordance with the invention are also known per se and are described in detail in the aforementioned textbooks. They embrace the following procedures:

a) Transformation of a suitable host organism, advantageously E.coli, with an expression vector in which an aforementioned hybrid gene is operatively bonded to an expression control sequence;

b) cultivation of the thus-obtained host organism under suitable growth conditions; and

c) extraction and isolation of the desired fusion protein from the host organism.

[0015] As host organisms there come into consideration gram-negative and gram-positive bacteria, for example E. coli and B.subtilis strains. E.coli strain M15 is an especially preferred host organism of the present invention. Apart from the above-mentioned E.coli strain there can, however, also be used other generally accessible E. coli strains, for example E. coli 294 (ATCC No. 3144), E.coli RR1 (ATCC No. 31343) and E. coli W3110 (ATCC No. 27325).

[0016] The fusion proteins in accordance with the present invention allow the selected cleavage at a specific chemical or enzymatic cleavage site without affecting the desired hydrophobic polypeptide, protein or peptide. Diffusion of the desired hydrophobic polypeptide, protein or peptide into the solid phase of a hydrophobic matrix column enables to orient the fusion proteins in accordance with the present invention so as to hide the desired hydrophobic polypeptide. protein or peptide. The bulky hydrophilic peptide on the other hand exposes the selected cleavage site to the mobile aqueous phase. This allows one to remove the bulky hydrophilic peptide by selected cleavage leaving only the desired hydrophobic polypeptide, protein or peptide bound to the column. The desired hydrophobic polypeptide, protein or peptide can then be eluted by addition of organic solvents.

[0017] Hence, the present invention also provides a process allowing the production and purification of a desired hydrophobic polypeptide protein or peptide, which process comprises the steps of:

a) passing an aqueous solution containing a fusion protein in accordance with the present invention through a hydrophobic matrix column,

b) flushing the column with a solution containing a cleavage reagent or an enzyme, and

c) removing the resulting desired hydrophobic polypeptide, protein or peptide with an aqueous water miscible solvent.

[0018] As hydrophobic matrix columns there come into consideration cyanopropyl, cyclohexyl, phenyl, octyl or octadecyl group bonded silica matrix columns. In the preferred practice of the invention RP - 18 (octadecyl bound silica microparticle column) under reversed phase high performance liquid chromatography (HPLC) conditions is used.

[0019] Prior to the loading with the fusion protein in accordance with the invention, the hydrophobic matrix column is conveniently equilibrated with an aqueous buffer. The equilibration buffer can contain a denaturing agent or a chaotropic agent, for example guanidine-HCL, urea or a detergent, e.g. Triton. The addition of such a denaturing agent, chaotropic agent or detergent permits problem-free operations even with fusion proteins in accordance with the invention which are extremely difficult to solublize in aqueous solution.

[0020] The fusion protein in accordance with the present invention is applied onto the hydrophobic matrix column in aqueous buffer which can also contain a denaturing agent or a detergent. for example guanidine-HCL, urea or Triton.

[0021] Cleavage is performed by flushing the column with an aqueous buffer containing a cleavage reagent or an enzyme. The optimal buffer composition depends on the cleavage reagent or enzyme used and is conveniently determined on a case-by-case basis.

[0022] The elution of the desired hydrophobic polypeptides, proteins or peptides can be carried out using a gradient of an aqueous water miscible solvent. Suitable water miscible solvents for this purpose include alkanols such as n-propanol, 2-propanol, ethanol, methanol, tert- butanol or cyclic ethers such as dioxane. The optimal elution conditions depend on the desired hydrophobic polypeptide, protein or peptide to be purified, the hydrophobic matrix, the column dimensions etc. and are conveniently determined on a case-by-case basis.

[0023] The aforementioned process allowing the production and purification of a desired hydrophobic polypeptide protein or peptide can also be carried out batch-wise. The fusion protein in accordance with the present invention is then absorbed to a hydrophobic matrix in aqueous buffer. Cleavage is performed by incubating the hydrophobic matrix with an aqueous buffer containing the cleavage reagent or the enzyme. The desired hydrophobic polypeptide can be obtained by incubating the hydrophobic matrix with the aqueous water miscible solvent after removal of the cleavage reagent or enzyme and the bulky hydrophilic peptides.

[0024] The novel process allowing production and purification of a desired hydrophobic polypeptide, protein or peptide is also employed to purify HIV-1 envelope peptide with Seq ID No:11 to homogeneity using analogous conditions as those used in reference Example 1.

[0025] The peptide with Seq ID No:11 obtained by the novel process may be used in diagnosing HIV infections.

[0026] These Examples can be understood better when they are read in conjunction with the accompanying Figures. The following symbols appear in these Figures:

'N250PSN250P29' represents the regulatable promoter/operator element N25OPSN25OP29, 'RBSII' represents the synthetic ribosomal binding site RBSII; '[His]6', '[NANP]19' and 'amy' represent the genes encoding the 6xHis-NANP-amyloid fusion proteins of this invention; 'bla', 'cat', 'lacI' and 'neo' represent the genes for beta-lactamase, chloramphenicol acetyltransferase, lac repressor and neomycin phosphotransferase, respectively; 'to', 'TE' and T1' represent transcriptional terminators t_o of phage lambda, TE of phage T7 and T1 of the E. coli rrnB operon; 'repl.' represents the replication regions of plasmids pBR322 and pREP4.

Figure 1
is a schematic drawing of the plasmid pREP4.

Figure 2
is a schematic drawing of the plasmid p6xHis-NANP-Met-Amy.

Figure 3
displays that part of the nucleotide sequence of plasmid p6xHis-NANP-Met-Amy (Seq ID No: 7) which encodes the fusion protein 6xHis-NANP-Met-Amy (Seq ID No: 8). In this sequence, the recognition sequences of some of the restriction enzymes depicted in Figure 2 are indicated. The amino acid sequence shown represents in the three letter code the sequence of the fusion protein 6xHis-NANP-Met-Amy, amino acids corresponding to the bA4-peptide are numbered.

Figure 4
is a schematic drawing of the plasmid pB/E1-6xHis-NANP-Met-huAmy.

Figure 5
displays that part of the nucleotide sequence of plasmid pB/E1-6xHis-NANP-Met-huAmy (Seq ID No: 9) which encodes the fusion protein 6xHis-NANP-Met-huAmy (Seq ID No: 10). In this sequence, the recognition sequences of some of the restriction enzymes depicted in Figure 4 are indicated. The amino acid sequence shown represents in the three letter code the sequence of the fusion protein 6xHis-NANP-Met-huAmy, amino acids corresponding to the bA4-peptide are numbered. The lower part of the Figure displays the nucleotide sequences and the encoded amino acids by which plasmid pB/E1-6xHis-NANP-Met-huAmy[M35E], which

encodes the fusion protein 6xHis-NANP-Met-huAmy[M35E], plasmid pB/E1-6xHis-NANP-Met-huAmy[M35L], which encodes the fusion protein 6xHis-NANP-Met-huAmy[M35L], plasmid pB/E1-6xHis-NANP-Met-huAmy [M35Q], which encodes the fusion protein 6xHis-NANP-Met-huAmy[M35Q], and plasmid pB/E1-6xHis-NANP-Met-huAmy[M35S], which encodes the fusion protein 6xHis-NANP-Met-huAmy[M35S], differ from plasmid pB/E1-6xHis-NANP-Met-huAmy.

Figure 6
gives the results of bA4 analysis by non-denaturing agarose gel electropheresis. This analysis was performed with "Serum Protein Electrophoresis system Paragon" from Beckman according to the recommendations of the supplier. 8 µg (in 2 µl H_2O) of peptide each were applied. Staining was with Coomassie brilliant blue for 3 h and destaining with 10% acetic acid, 45% methanole and 45% water. Samples were prepared in destilled water immediately before the experiment (f) or allowed to age for 2 days (a). The samples tested are given in the list below.

gel a		gel b	
lanel	human bA4 wash 1 f	lane 1	M35S bA4 wash 1 f
lane 2	human bA4 wash 2 f	lane 2	M35S bA4 wash 2 f
lane 3	human bA4 wash 3 f	lane 3	M35S bA4 wash 3 f
lane 4	human bA4 wash 4 f	lane 4	M35S bA4 wash 4 f
lane 5	human bA4 wash 4 a	lane 5	M35S bA4 wash 4 a
lane 6	Standard BSA 2 mg	lane 6	Standard BSA 10 mg
lane 7	M35L bA4 wash 1 f	lane 7	M35Q bA4 wash 1 f
lane 8	M35L bA4 wash 2 f	lane 8	M35Q bA4 wash 2 f
lane 9	M35L bA4 wash 3 f	lane 9	M35Q bA4 wash 3 f
lane 10	M35L bA4 wash 3a	lane 10	rat bA4
gel c			

gives the results of a comparison of the point mutants M35Q bA4 and M35E bA4. The point mutant M35E bA4 contains an extra negative charge which results in a higher mobility on the Beckman gel (lane 8: M35E bA4; lane 9: M35Q bA4). When M35E bA4 is mixed with M35Q bA4 after cleavage (lane 1) or when the fusion protein containing M35Q bA4 is mixed with the fusion protein containing M35E bA4 and cleaved according to the process of the present invention (lanes 3, 4, 5 and 6 (wash 1, 2, 3 and 4)) a clear separation is observed indicating the presence of monomeric bA4.

Figure 7
gives the results of bA4 chromatography on a size-exclusion column.

[0027] Fractions of 250 ml each were collected and analyzed by non-denaturing electrophoresis. Lane 1: bA4 applied onto the column. Lanes 2 to 10: peak fractions.

[0028] The marker proteins used to determine the size of bA4 were :

serum albumin	(MW = 65000 ; retention time = 19.89 min) ,
ovalbumine	(MW = 45000 ; retention time = 20.10 min) ,
lactalbumin	(MW = 14200 ; retention time = 21.43 min) and
insulin	(MW = 5734 ; retention time = 21.68 min) .

[0029] The retention time of 22.54 min. points to a monomer with a MW of about 4500 daltons. This is in agreement with light-scattering data and ultracentrifugation experiments by the Yphantis method (D.A. Yphantis. Annals of the N. Y. Acad.Sci. 88, 586-601 [1960]).

Example 1

Expression plasmid used for the preparation of the fusion protein 6xHis-NANP-Met-Amy

[0030] The expression plasmid p6xHis-NANP-Met-Amy (see Figures 2 and 3) was used for the preparation of the

fusion protein 6xHis-NANP-Met-Amy. E. coli M15 cells transformed with plasmids pREP4 and p6xHis-NANP-Met-Amy were deposited in accordance with the Budapest Treaty at the Deutsche Sammlung von Mikroorganismen (DSM) in Braunschweig. BRD, on May 18, 1993, under the accession number DSM 8310.

Example 2

Expression plasmid used for the preparation of the fusion protein 6xHis-NANP-Met-huAmy

[0031] The expression plasmid pB/E1-6xHis-NANP-Met-huAmy (see Figures 4 and 5) was used for the preparation of the fusion protein 6xHis-NANP-Met-huAmy. E. coli M15 cells transformed with plasmids pREP4 and pB/E1-6xHis-NANP-Met-huAmy were deposited in accordance with the Budapest Treaty at the Deutsche Sammlung von Mikroorganismen (DSM) in Braunschweig, BRD, on May 18, 1993, under the accession number DSM 8311.

Expression plasmids used for the preparation of the fusion proteins 6xHis-NANP-Met-huAmy[M35L], 6xHis-NANP-Met-huAmy[M35Q], 6xHis-NANP-Met-huAmy[M35S], and 6xHis-NANP-Met-huAmy [M35E]

[0032] The expression plasmids pB/E1-6xHis-NANP-Met-huAmy[M35L], pB/E1-6xHis-NANP-Met-huAmy[M35Q], pB/E1-6xHis-NANP-Met-huAmy[M35S] and pB/E1-6xHis-NANP-Met-huAmy[M35E], which differ from plasmid pB/E1-6xHis-NANP-Met-huAmy only in the nucleotides encoding amino acid 35 of the bA4 amyloid peptide (see Figure 5), were used for the preparation of the fusion proteins 6xHis-NANP-Met-huAmy[M35L],6xHis-NANP-Met-huAmy [M35Q], 6xHis-NANP-Met-huAmy[M35S] and 6xHis-NANP-Met-huAmy[M35E], respectively. E. coli M15 cells transformed with plasmids pREP4 and pB/E1-6xHis-NANP-Met-huAmy[M35L], pREP4 and pB/E1-6xHis-NANP-Met-huAmy [M35Q], pREP4 and pB/E1-6xHis-NANP-Met-huAmy[M35S] and pREP4 and pB/E1-6xHis-NANP-Met-huAmy[M35E]], respectively, were deposited in accordance with the Budapest Treaty at the Deutsche Sammlung von Mikroorganismen (DSM) in Braunschweig, BRD, on May 18, 1993, under the accession numbers DSM 8313, DSM 8314, DSM 8315 and DSM 8312, respectively.

Example 4

Fermentation and purification of fusion proteins

Fermentation

[0033] Plasmids p6xHis-NANP-Met-Amy, pB/E1-6xHis-NANP-Met-huAmy, pB/E1-6xHis-NANP-Met-huAmy [M35 L], pB/E1-6xHis-NANP-Met-hu Amy[M35Q] pB/E1-6xHis-NANP-Met-huAmy [M35S] and pB/E1-6xHis-NANP-Met-huAmy [M35E], respectively were transformed into E. coli M15 cells already containing plasmid pREP4 by standard methods (Sambrook et al., supra). Transformed cells were grown at 37°C in a 100 l fermenter in Super medium [Stüber et al., Immunological Methods, eds. Lefkovits and Pernis, Academic Press, Inc.. Vol.IV, 121-152 [1990]] containing 100 mg/l ampicillin and 25 mg/l kanamycin. At an optical density at 600 nm of about 1.0 IPTG was added to a final concentration of 2 mM. After an additional 3 hrs at 37°C the cells were harvested by centrifugation. In a typical fermentation run a biomass of approximately 500 g containing at least 3 g of the recombinant fusion protein was obtained.

Purification

[0034] 2.5 l of 6 M guanidine·HCl containing 0.1 M di-sodium hydrogen phosphate, pH 8 were added to the cells and stirred for 24 hours. Crude cell debris were removed by centrifugation and the supernatant was then further clarified by cross-flow filtration using a 0.3 mm membrane. The protein contained in the filtrate was then adsorbed to a Ni-NTA column (5 cm x 24 cm, flow 20 ml/min). Contaminating *E.coli* proteins were removed by washing first with 8M urea, pH 7.5. Elution was performed with 8 M urea, pH 4 . The chromatogram was monitored by SDS-PAGE and fractions containing fusion protein were pooled. A small aliquot of that pool was mixed with EDTA and desalted by dialysing against water, lyophilized and then analysed by electron spray mass-spectrometry (Table 1).

Table 1 :

Characterization of the fusion proteins			
Fusion Protein	Purification yield per 100 l fermenter [gram]	Theoretical mass	Average mass by electron spray MS
Human WT	3.0	13817	13820
Human Mut. M35S	5.5	13773	13776
Human Mut. M35L	5.7	13799	13802
Human Mut M35Q	4.5	13814	13817
Human Mut. M35E	6.0	13813	not tested
Rat WT	6.0	13724	13728
WT = wild-type, Mut = mutant, e.g. M35S (Met at position 35 is mutated to Ser)			

Reference Example 1

Cleavage of the fusion proteins to yield 1-42 b-amyloid peptides

[0035] A semipreparative RP-18 HPLC (Vidac, specification 218TP 152010, 250 mm x 10 mm) column was first equilibrated with 8 M urea, pH 4, at a flow rate of 2 ml/min. Then an aliquot of the NTA-eluate containing 400 mg fusion protein in 8 M urea, pH 4, was pumped onto the column at a flow rate of 1 ml/min. Then the column was washed with 8 M urea, pH 4, at a flow rate of 2 ml/min. The urea was washed out by water at a flow rate of 2 ml/min until baseline adsorbance at the column outlet was reached. Cleavage was performed by flushing the column with 45 mg/ml CNBr in a solution composed of 20% ethanol, 40% formic acid and 40% water for 24 hrs at 22 °C at a flow rate of 0.5 ml/min. The column was then flushed with 0.1 M EDTA at a flow rate of 2 ml/min and CNBr together with liberated MRG-SHHHHHHGS-$(NANP)_{19}$-RSM was washed out with 0.05% trifluoro acetic acid at a flow rate of 2.0 ml/min. 1-42 residue b-amyloid peptide was eluted at a flow rate of 2 ml/min using the following ethanol gradient given by the time points : (min/% ethanol) 0/0, 40/40, 45/50, 50/65, 55/100, 60/100, 65/0

[0036] A broad peak containing the b-amyloid peptide emerges between 45 and 60 min. Critical for the peptide to be monomeric was the immediate dilution with distilled water (e.g. by dropping the eluate into a stirred beaker containing 200 ml H_2O), and immediate lyophilization. The resulting powder was named W1. Since a considerable amount of b-amyloid peptide remained on the column, the elution was repeated three times using the above mentioned protocol, giving rise to samples W2, W3 and W4. These samples were tested for purity (Table 2) and the amont of monomeric bA4 [Figure 6]. The correct chemical structure of the peptides was verified by electron spray mass-spectrometry (Table 3), by amino acid analysis and by amino terminal sequencing (Table 3).

Table 2 :

Production of 1-42 b-amyloid peptides				
Peptide	Number of tests	Wash Number	Mass produced (mg)	Purity (%)
rat WT (Seq ID No:2)	4	1	100 ± 14	90 ± 2
		2	28 ± 8	89 ± 2
		3/4	9 ± 3	90 ± 1
human WT (Seq ID No:1)	3	1	96 ± 23	79 ± 9
		2	52 ± 10	73 ± 5
		3	32 ± 9	78 ± 4
		4	15 ± 8	86 ± 5

Table 2 : (continued)

Production of 1-42 b-amyloid peptides				
Peptide	Number of tests	Wash Number	Mass produced (mg)	Purity (%)
human M35S (Seq ID No:4)	3	1	71 ± 13	80 ± 8
		2	31 ± 5	74 ± 2
		3	18 ± 7	83 ± 2
		4	2 ± 7	88 ± 2
human M35L (Seq ID No:3)	3	1	59 ±15	72 ± 8
		2	23 ± 9	71 ± 4
		3	8 ± 3	78 ± 3
human M35Q (Seq ID No:5)	4	1	37 ± 4	70 ± 2
		2	10 ± 5	82 ± 3
human M35E (Seq D No:6)	1	1	95	not tested
		2	68	not tested
		3	46	not tested
		4	31	not tested

[0037] Purity is based on amino acid analysis, the content of fusion protein is detected by the quotient of Asp to Glu. Pure 1-42 b-amyloid peptide has 4 Glu and 4 Asp residues, pure fusion protein has a ratio of 42 Asp to 4 Glu.

Table 3 :

Identification of 1-42 b-amyloid peptides			
Sample	Mass Spectrometry		Edman degradation
	Theory	Found	10-15 cycles
human WT	4515.1	4531**	DAEFRHDSGYEVHHQ
human M35S	4471.0	4472	DAEFRHDSGY
human M35L	4497.1	4498	DAEFRHDSGY
human M35Q	4512.0	4512	not tested
human M35E	4513.0	4512	not tested
rat WT	4417.0	4435**	DAEFGHDSGF

** The methionines of human WT and rat WT were transformed during the cleavage procedure to methionine sulfoxide.

Reference Example 2

Purification of monomeric 1-42 bA4

Small scale method

[0038] Samples containing 1 mg of bA4 were applied to a LKB UltroPac HPLC column (diameter: 7.5 mm, length : 600 mm, flow 0.5 ml/min, buffer: 12mM Tris(hydroxymethyl)aminomethan containing 200 mM glycine, pH 7.8). bA4 emerged in a sharp peak, and the peak fractions contained the Coomassie blue band on the agarose electrophoresis gel shown in Figure 7. Light-scattering experiments gave no evidence of high molecular weight forms (aggregates or fibres) and calibration standards which were chromatographed under identical conditions pointed to a molecular weight of about 4500, indicating that bA4 was present as a monomer.

Preparative scale method

[0039] The "Continuous Elution Electrophoresis system Model 491 Prep Cell" from BioRad was used. The non-denaturating discontinuous acrylamide gel was composed of a 4 % acrylamide / 2.7 % N,N'-methylene-bis-acrylamide separating gel in 0.375M Tris(hydroxymethyl) amino-methan (pH8.8) with a length of 3 cm and a diameter of 37 mm, and equipped with a cooling tube of 20 mm diameter.

[0040] The stacking gel was composed of 4 % acrylamide / 2.7 % N,N'-methylene-bis-acrylamide in 0,125 M Tris (hydroxymethyl)amino-methan (pH 6.8) and had a length of 2 cm. The running buffer was 25 mM Tris(hydroxymethyl)aminomethan / 0.2 M glycine, pH 8.3 . For elution the same buffer was used with a flow rate of 0.75 ml/minutes. 45 mg of bA4 peptide, generally from wash 2, 3 or 4, were dissolved in 7 ml H_2O and 1 ml glycerole. Electrophoresis was performed at 12 Watt (constant, limits at 500 V and 40 mA), a typical run requiring 4 hours. The results of a typical run are shown in Figure 7.

Example 5

Purification of HIV-1 envelope peptide

[0041] 400 mg fusion protein MRGS $(H)_6$ GS $(NANP)_{19}$ RSM RILA VERYLKDQQLLGIWGCSGKLICTTAVPWNAS (Seq ID No:12) prepared and purified by the same methods as those described for the preparation and purification of the fusion proteins containing monomeric 1-42 bA4 (Examples 1 - 4) were loaded onto a Vydac RP 18 column (specification 218 TP 152010, 250 mm x 10 mm) and cleaved using the same conditions as those described for the cleavage of the 1 - 42 bA4 fusion proteins (Reference Example 1). Approximately 20 - 30 mg of lyophilised powder were obtained. This powder was analysed by electron spray analysis. A peak of 3902 ± 2 Da corresponding to the HIV-1 envelope peptide (Seq ID No:11) was detected.

SEQUENCE LISTING

[0042]

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: F. HOFFMANN-LA ROCHE AG
(B) STREET: Grenzacherstrasse 124
(C) CITY: Basle
(D) STATE: BS
(E) COUNTRY: Switzerland
(F) POSTAL CODE (ZIP): CH-4002
(G) TELEPHONE: 061 - 688 42 56
(H) TELEFAX: 061 - 688 13 95
(I) TELEX: 962292/965542 hlr ch

(ii) TITLE OF INVENTION: Process for producing hydrophobic polypeptides, proteins or peptides

(iii) NUMBER OF SEQUENCES: 12

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Met Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Asp Ala Glu Phe Gly His Asp Ser Gly Phe Glu Val Arg His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Met Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Leu Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Ser Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Gln Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Glu Val Gly Gly Val Val Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 520 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 115..516
(D) OTHER INFORMATION: /product= "Amyloid Protein AA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA
60

ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACT ATG
117

Met
1

AGA GGA TCG CAT CAC CAT CAC CAT CAC GGA TCT AAC GCG AAC CCG AAC 165
Arg Gly Ser His His His His His His Gly Ser Asn Ala Asn Pro Asn
5 10 15

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
213
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
20 25 30

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
261
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
35 40 45

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
309
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
50 55 60 65

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
357
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
70 75 80

GCG AAC CCG AAC GCG AAC CCG AGA TCT ATG GAT GCG GAG TTC GGA CAT 405
Ala Asn Pro Asn Ala Asn Pro Arg Ser Met Asp Ala Glu Phe Gly His
85 90 95

GAT TCA GGC TTC GAA GTC CGC CAT CAA AAA CTG GTG TTC TTT GCA GAA 453
Asp Ser Gly Phe Glu Val Arg His Gln Lys Leu Val Phe Phe Ala Glu
100 105 110

GAT GTG GGT TCA AAC AAA GGT GCC ATC ATT GGA CTC ATG GTG GGT GGC 501

Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly
115 120 125

GTT GTC ATA GCA TAAGCTT 523
Val Val Ile Ala
130

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Arg Gly Ser His His His His His His Gly Ser Asn Ala Asn Pro
1 5 10 15

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
20 25 30

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
35 40 45

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
50 55 60

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
65 70 75 80

Asn Ala Asn Pro Asn Ala Asn Pro Arg Ser Met Asp Ala Glu Phe Gly
85 90 95

His Asp Ser Gly Phe Glu Val Arg His Gln Lys Leu Val Phe Phe Ala
100 105 110

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
115 120 125

Gly Val Val Ile Ala
130

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 520 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 115..516

(D) OTHER INFORMATION: /product= "Amyloid Protein AA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA 60

ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACT ATG 117

Met

1

AGA GGA TCG CAT CAC CAT CAC CAT CAC GGA TCT AAC GCG AAC CCG AAC 165

Arg Gly Ser His His His His His His Gly Ser Asn Ala Asn Pro Asn

5 10 15

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC 213

Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn

20 25 30

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC 261

Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
35 40 45

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
309
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
50 55 60 65

GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC GCG AAC CCG AAC
357
Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn
70 75 80

GCG AAC CCG AAC GCG AAC CCG AGA TCT ATG GAT GCG GAG TTC CGT CAT 405
Ala Asn Pro Asn Ala Asn Pro Arg Ser Met Asp Ala Glu Phe Arg His
85 90 95

GAT TCA GGC TAT GAA GTC CAC CAT CAA AAA CTG GTG TTC TTT GCA GAA 453
Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu
100 105 110

GAT GTG GGT TCA AAC AAA GGT GCC ATC ATT GGA CTC ATG GTG GGT GGC 501
Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly
115 120 125

GTT GTC ATA GCA TAAGCTT 520
Val Val Ile Ala
130

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Arg Gly Ser His His His His His His Gly Ser Asn Ala Asn Pro
1 5 10 15

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
20 25 30

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
35 40 45

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
50 55 60

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
65 70 75 80

Asn Ala Asn Pro Asn Ala Asn Pro Arg Ser Met Asp Ala Glu Phe Arg
85 90 95

His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala
100 105 110

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
115 120 125

Gly Val Val Ile Ala
130

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
1 5 10 15

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp
20 25 30

Asn Ala Ser
35

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 126 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Arg Gly Ser His His His His His His Gly Ser Asn Ala Asn Pro
1 5 10 15

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
20 25 30

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
35 40 45

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
50 55 60

Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro Asn Ala Asn Pro
65 70 75 80

Asn Ala Asn Pro Asn Ala Asn Pro Arg Ser Met Arg Ile Leu Ala Val
85 90 95

Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser
100 105 110

Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser
115 120 125

Claims

1. A fusion protein of the formula:

A - B - C

wherein A is a bulky hydrophilic peptide wherein the bulky hydrophilic peptide has a peptide sequence of the formula

(NANP)x

wherein x is 10-40, preferably 19, B is a selected cleavage site and C is a desired hydrophobic polypeptide, protein or peptide.

2. A fusion protein in accordance with claim 1, wherein the selected cleavage site is a chemical cleavage site, preferably a methionine residue which site is specifically cleaved by cyanogen bromide.

3. A fusion protein in accordance with claim 1 or 2, wherein the desired hydrophobic peptide is the HIV-1 envelope peptide with [Seq ID No: 11].

4. Genes which code for a fusion protein in accordance with any one of claims 1-3.

5. Expression vectors in which a gene in accordance with claim 4 is operatively linked to an expression control sequence.

6. A bacterium transformed with an expression vector in accordance with claim 5.

7. A bacterium as claimed in claim 6 which is E.coli.

8. A process for production and purification of a desired hydrophobic polypeptide, protein or peptide, which process comprises the steps of :

   a) passing an aqueous solution containing fusion protein in accordance with any one of claims 1-3 through a hydrophobic matrix column.

b) flushing the column with a solution containing a cleavage reagent or an enzyme, and

c) removing the resulting desired hydrophobic polypeptide, protein or peptide with an aqueous water miscible solvent.

9. A process in accordance with claim 8, wherein the hydrophobic matrix column is an octadecyl bound silica microparticle column.

10. A process in accordance with claim 8 or 9, wherein the cleavage reagent is cyanogen bromide.

11. A process in accordance with claims 8 - 10 wherein the desired hydrophobic peptide is the HIV-1 envelope peptide with [Seq ID No. 11].

12. The use of a fusion protein in accordance with claims 1 - 3 for the production and purification of the desired hydrophobic polypeptide, protein or peptide C.

Patentansprüche

1. Fusionsprotein der Formel A-B-C, worin A ein sperriges hydrophiles Peptid ist, wobei das sperrige hydrophile Peptid eine Peptidsequenz der Formel $(NANP)_x$ hat, worin x 10 bis 40, bevorzugt 19 ist, B eine ausgewählte Schnittstelle ist und C ein gewünschtes hydrophobes Polypeptid, Protein oder Peptid ist.

2. Fusionsprotein nach Anspruch 1, wobei die ausgewählte Schnittstelle eine chemische Schnittstelle ist, bevorzugt ein Methioninrest, die spezifisch durch Bromcyan gespalten wird.

3. Fusionsprotein nach Anspruch 1 oder Anspruch 2, wobei das gewünschte hydrophobe Peptid das HIV-1-Envelope-Peptid mit [SEQ ID Nr. 11] ist.

4. Gene, die ein Fusionsprotein nach einem der Ansprüche 1 bis 3 codieren.

5. Expressionsvektoren, bei denen ein Gen gemäß Anspruch 4 operativ mit einer Expressionskontrollsequenz verbunden ist.

6. Bakterium, das mit einem Expressionsvektor nach Anspruch 5 transformiert ist.

7. Bakterium nach Anspruch 6, das *E. coli* ist.

8. Verfahren zur Herstellung und Reinigung eines gewünschten hydrophoben Polypeptids, Proteins oder Peptids, wobei das Verfahren die Stufen umfasst, dass:

a) eine wässrige Lösung, die ein Fusionsprotein nach einem der Ansprüche 1 bis 3 enthält, über eine Säule mit einer hydrophoben Matrix geleitet wird,
b) die Säule mit einer Lösung gespült wird, die ein Spaltreagenz oder ein Enzym enthält und
c) das entstehende gewünschte hydrophobe Polypeptid, Protein oder Peptid mit einem wässrigen wassermischbaren Lösungsmittel entfernt wird.

9. Verfahren nach Anspruch 8, wobei die Säule mit hydrophober Matrix eine Siliciumdioxidmikroteilchensäule mit gebundenen Octadecylgruppen ist.

10. Verfahren nach Anspruch 8 oder Anspruch 9, wobei das Spaltreagenz Bromcyan ist.

11. Verfahren nach einem der Ansprüche 8 bis 10, wobei das gewünschte hydrophobe Peptid HIV-1-Envelope-Peptid mit [SEQ ID Nr. 11] ist.

12. Verwendung eines Fusionsproteins nach einem der Ansprüche 1 bis 3 zur Herstellung und Reinigung des gewünschten hydrophoben Polypeptids, Proteins oder Peptids C.

Revendications

1. Protéine de fusion de formule

$$A\text{-}B\text{-}C$$

dans laquelle
A est un peptide hydrophile volumineux, le peptide hydrophile volumineux ayant une séquence peptidique de formule

$$(NANP)_x$$

dans laquelle x est un nombre de 10 à 40, de préférence 19,
B est un site de clivage choisi et
C est un polypeptide, une protéine ou un peptide hydrophobe désiré(e).

2. Protéine de fusion selon la revendication 1, dans laquelle le site de clivage choisi est un site de clivage chimique, de préférence un résidu de méthionine, ce site étant clivé de façon spécifique par du bromure de cyanogène.

3. Protéine de fusion selon la revendication 1 ou 2, dans laquelle le peptide hydrophobe désiré est le peptide d'enveloppe de VIH-1 ayant la séquence [SEQ ID N° 11].

4. Gènes qui codent pour une protéine de fusion selon l'une quelconque des revendications 1 à 3.

5. Vecteurs d'expression dans lesquels un gène selon la revendication 4 est lié de façon opérationnelle à une séquence de contrôle de l'expression.

6. Bactérie transformée avec un vecteur d'expression selon la revendication 5.

7. Bactérie selon la revendication 6, qui est *E. coli.*

8. Procédé de production et de purification d'un polypeptide, d'une protéine ou d'un peptide hydrophobe désiré(e), ce procédé comprenant les étapes selon lesquelles:

   (a) on fait passer une solution aqueuse contenant une protéine de fusion selon l'une quelconque des revendications 1 à 3 à travers une colonne de matrice hydrophobe,
   (b) on rince la colonne avec une solution contenant un réactif de clivage ou une enzyme, et
   (c) on sépare le polypeptide, la protéine ou le peptide hydrophobe désiré(e) obtenu(e) avec un solvant aqueux miscible à l'eau.

9. Procédé selon la revendication 8, dans lequel la colonne de matrice hydrophobe est une colonne de microparticules de silice ayant des groupes octadécyle liés.

10. Procédé selon la revendication 8 ou 9, dans lequel le réactif de clivage est le bromure de cyanogène.

11. Procédé selon les revendications 8 à 10, dans lequel le peptide hydrophobe désiré est le peptide d'enveloppe de VIH-1 ayant la séquence [SEQ ID N° 11].

12. Utilisation d'une protéine de fusion selon les revendications 1 à 3 pour la production et la purification du polypeptide, de la protéine ou du peptide hydrophobe C désiré(e).

Figure 1

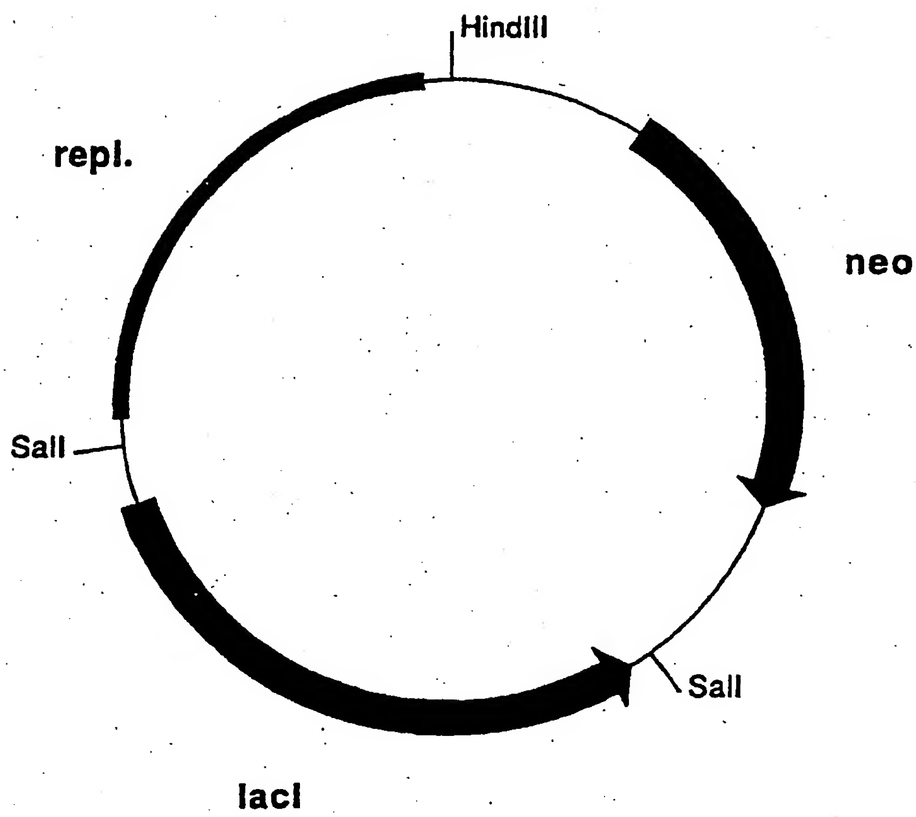


Figure 2

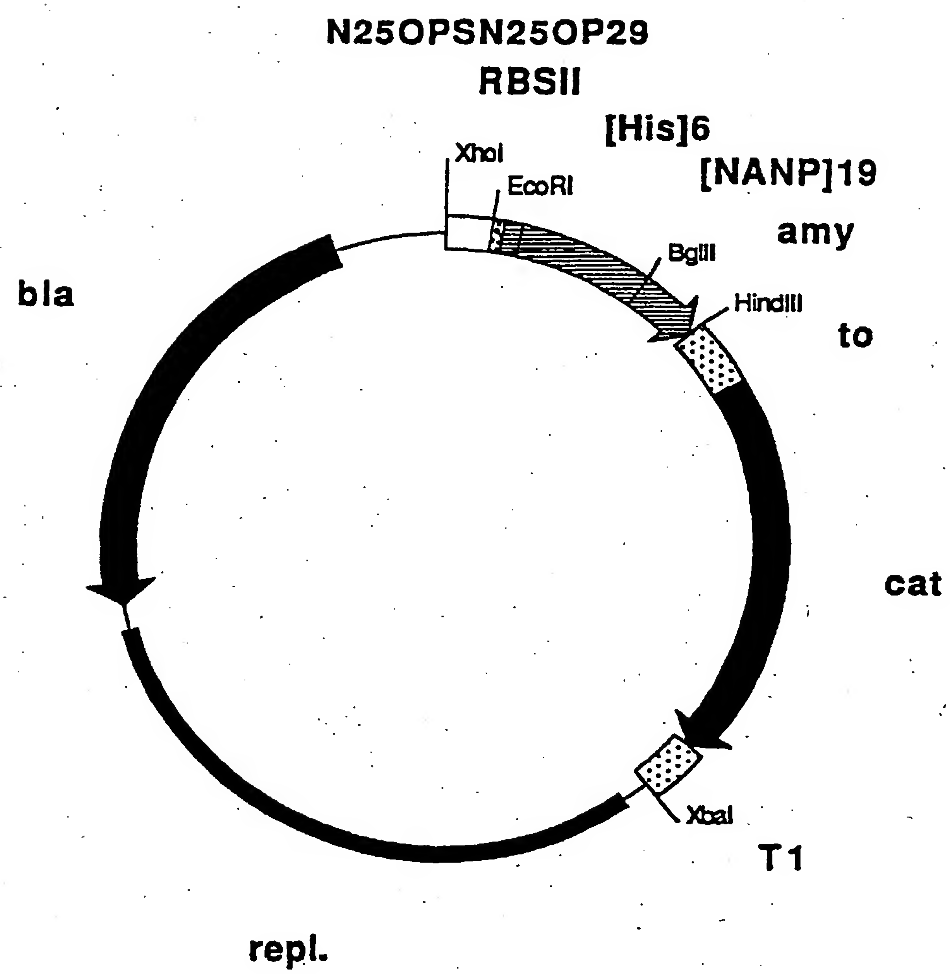

N25OPSN25OP29
RBSII
[His]6
[NANP]19
amy
XhoI
EcoRI
BglII
HindIII
to
bla
cat
XbaI
T1
repl.

Figure 3

```
        XhoI
  1 CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA

                                      EcoRI
 61 ATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGA
                                                           MetArg

121 GGATCGCATCACCATCACCATCACGGATCTAACGCGAACCCGAACGCGAACCCGAACGCG
    GlySerHisHisHisHisHisHisGlySerAsnAlaAsnProAsnAlaAsnProAsnAla

181 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

241 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

301 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

                      BglII
361 AACCCGAACGCGAACCCGAGATCTATGGATGCGGAGTTCGGACATGATTCAGGCTTCGAA
    AsnProAsnAlaAsnProArgSerMetAspAlaGluPheGlyHisAspSerGlyPheGlu
                               1                             11

421 GTCCGCCATCAAAAACTGGTGTTCTTTGCAGAAGATGTGGGTTCAAACAAAGGTGCCATC
    ValArgHisGlnLysLeuValPhePheAlaGluAspValGlySerAsnLysGlyAlaIle
                               21                            31

                                             HindIII
481 ATTGGACTCATGGTGGGTGGCGTTGTCATAGCATAAGCTT   520 (Seq. ID No:7)
    IleGlyLeuMetValGlyGlyValValIleAla              (Seq. ID No:8)
                               41
```

Figure 4

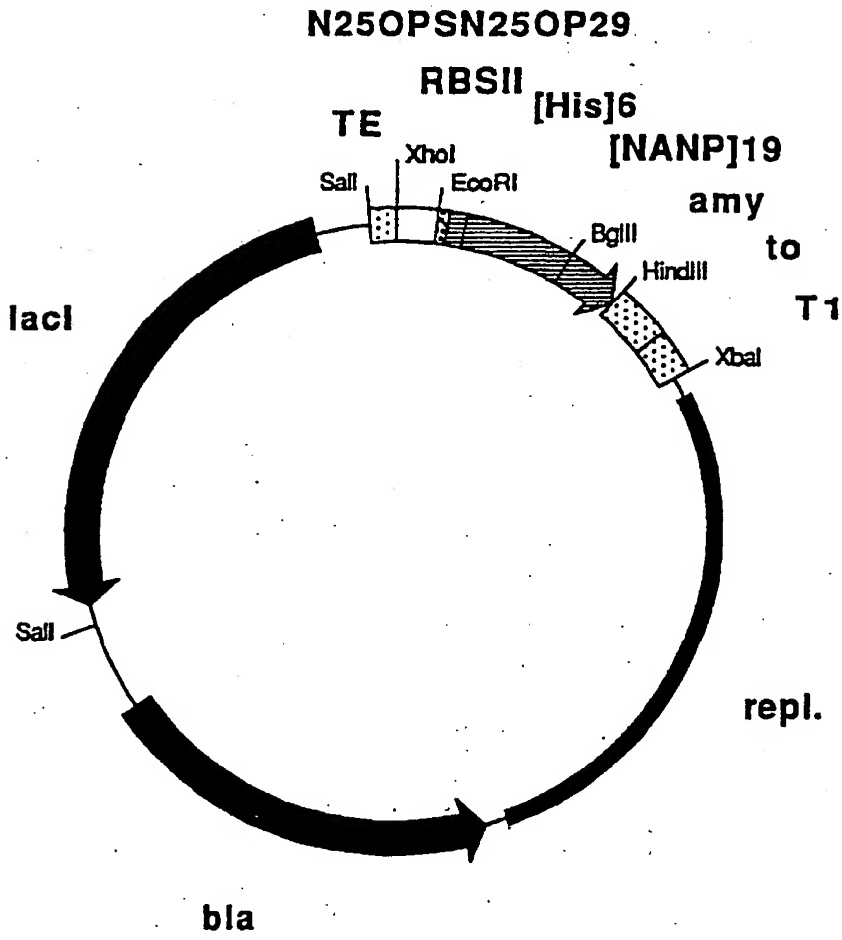

N25OPSN25OP29
RBSII
[His]6
TE
XhoI
[NANP]19
SalI
EcoRI
amy
BglII
to
HindIII
lacI
T1
XbaI
SalI
repl.
bla

Figure 5

```
  XhoI
1 CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA

                                               EcoRI
61 ATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGA
                                                          MetArg

121 GGATCGCATCACCATCACCATCACGGATCTAACGCGAACCCGAACGCGAACCCGAACGCG
    GlySerHisHisHisHisHisHisGlySerAsnAlaAsnProAsnAlaAsnProAsnAla

181 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

241 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

301 AACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCGAACCCGAACGCG
    AsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAlaAsnProAsnAla

                   BglII
361 AACCCGAACGCGAACCCGAGATCTATGGATGCGGAGTTCCGTCATGATTCAGGCTATGAA
    AsnProAsnAlaAsnProArgSerMetAspAlaGluPheArgHisAspSerGlyTyrGlu
                              1                             11

421 GTCCACCATCAAAAACTGGTGTTCTTTGCAGAAGATGTGGGTTCAAACAAAGGTGCCATC
    ValHisHisGlnLysLeuValPhePheAlaGluAspValGlySerAsnLysGlyAlaIle
                              21                            31

                                           HindIII
481 ATTGGACTCATGGTGGGTGGCGTTGTCATAGCATAAGCTT   520  (Seq.ID No: 9)
    IleGlyLeuMetValGlyGlyValValIleAla               (Seq.ID No.10)
          35                  41

[M35E]:  GAG      [M35L]:  CTG      [M35Q]:  CAG      [M35S]:  TCT
         Glu               Leu               Gln               Ser
         35                35                35                35
```

Figure 6

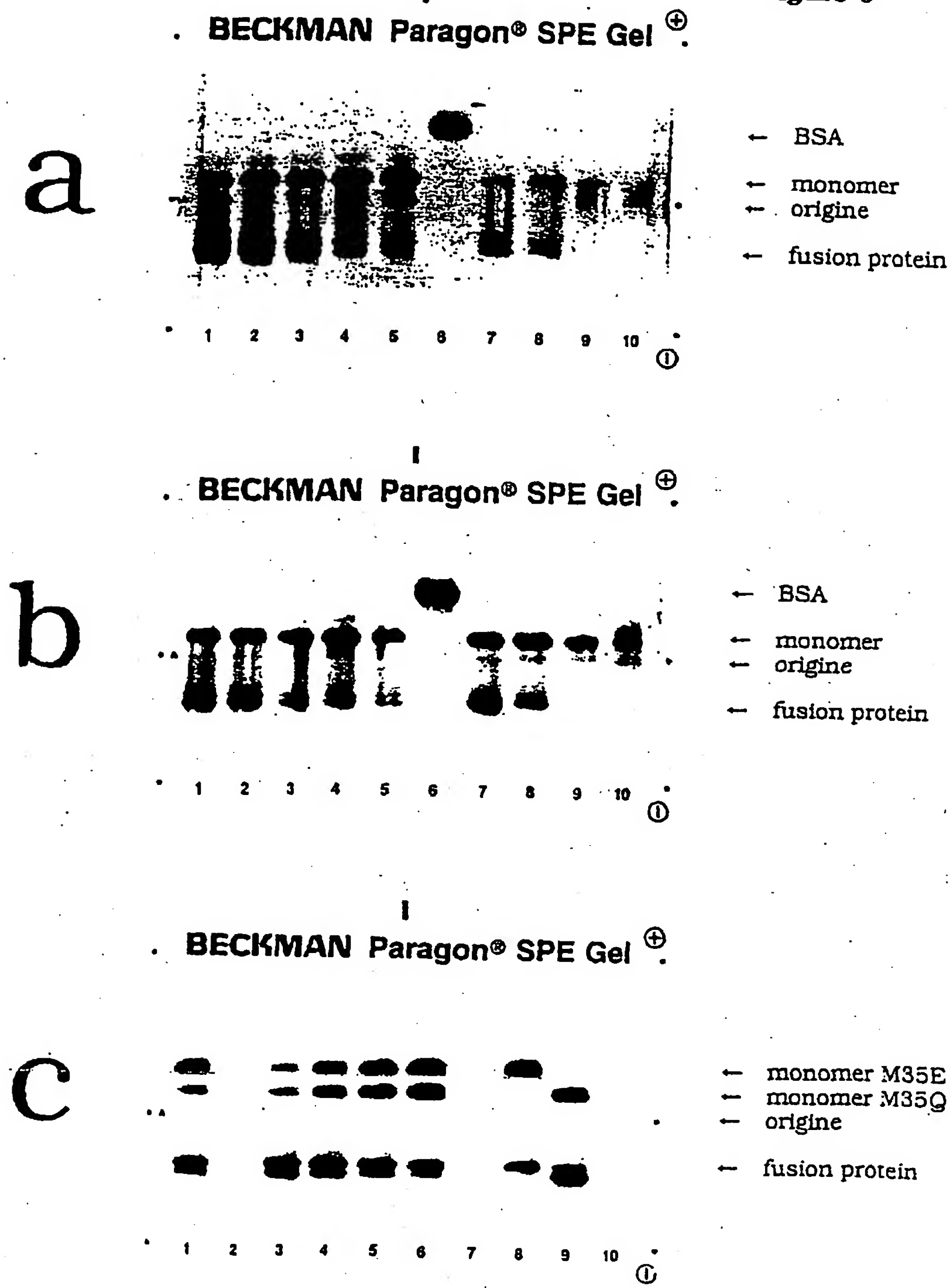
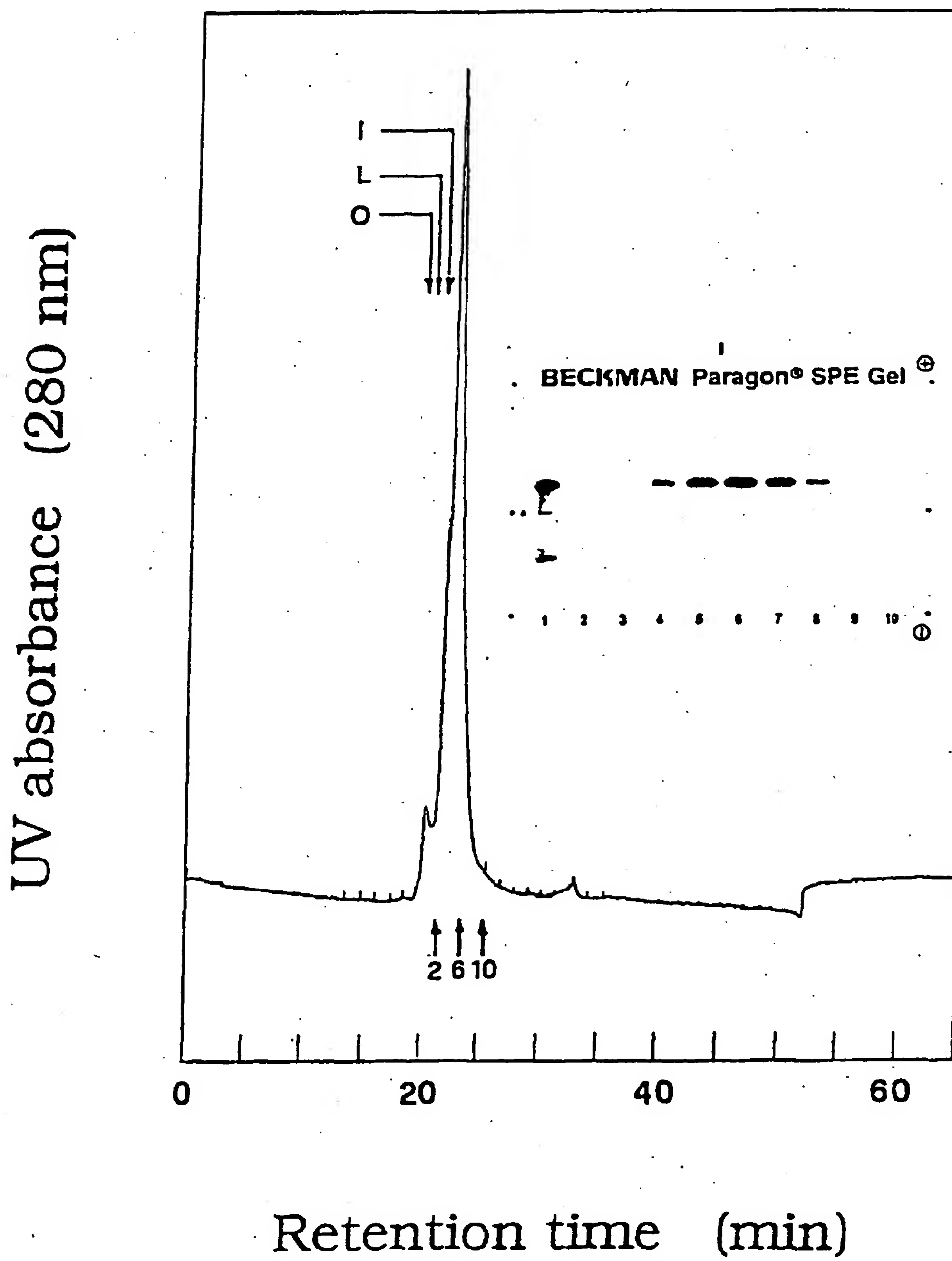

BECKMAN Paragon® SPE Gel
a
BSA
monomer
origine
fusion protein
1 2 3 4 5 6 7 8 9 10
BECKMAN Paragon® SPE Gel
b
BSA
monomer
origine
fusion protein
1 2 3 4 5 6 7 8 9 10
BECKMAN Paragon® SPE Gel
c
monomer M35E
monomer M35Q
origine
fusion protein
1 2 3 4 5 6 7 8 9 10

Figure 7


I
L
O
BECKMAN Paragon® SPE Gel
1 2 3 4 5 6 7 8 9 10
2 6 10
0
20
40
60
UV absorbance (280 nm)
Retention time (min)